



Veštačka inteligencija u službi forenzike glasa i govora

Mia M. Šešum*, Marina N. Šestić**, Maja P. Ivanović***

Univerzitet u Beogradu – Fakultet za specijalnu edukaciju i rehabilitaciju, Beograd, Srbija

Uvod: Veštačka inteligencija predstavlja vrhunac razvoja savremenih tehnologija. Programiran da pomaže ljudima u obavljanju operacija koje opterećuju ili čak i prevazilaze njihove mogućnosti, ovaj kompjuterski sistem pronalazi primenu u različitim područjima profesionalnog i svakodnevnog funkcionisanja. Jednu od oblasti u kojima resursi koje pruža veštačka inteligencija mogu biti dragoceni predstavljaju i forenzičke nauke. *Cilj:* Cilj rada je da se pomoću sistematizacije dostupne literature predstave mogućnosti veštačke inteligencije u kontekstu forenzike glasa i govora, zatim ograničenja, kao i potencijalne opasnosti od njene primene u sudskim postupcima. *Metode:* Literatura korišćena u radu obezbeđena je pretragom dostupnih baza podataka putem udaljenog pristupa, nakon čega se pristupilo sintezi relevantnih nalaza. *Rezultati:* Veštačka inteligencija je visokosofisticirano tehnološko dostignuće koje ima potencijal da u budućnosti transformiše različite forenzičke discipline. Zbog činjenice da je jezik jedan od najkompleksnijih društvenih fenomena, ali i zbog posledica koje bi mogle proisteći na osnovu analize u čiji način sprovođenja veštaci nemaju uvid, ova tehnologija nije uvrštena u zvanične i prihvaćene metode veštačenja glasa. *Zaključak:* Veštačka inteligencija može doprineti forenzičkim ispitivanjima glasa i govora, ali s obzirom na to da nedovoljna transparentnost ovih sistema nameće suštinska etička pitanja, njena primena za sada ostaje ograničena na operacije sadržane u pripremnom delu postupka veštačenja. Povećanje transparentnosti sistema zasnovanih na veštačkoj inteligenciji, uz podrazumevane konstantne napore ka poboljšanju otpornosti i preciznosti, doprineće rešavanju etičkih nedoumica i prihvatanju ove tehnologije kao legitimne u sudskim postupcima.

Ključne reči: forenzika, glas i govor, veštačka inteligencija, automatska identifikacija, biometrija

Korespondencija: Mia Šešum, gja982@gmail.com

* <https://orcid.org/0000-0001-5192-878X>

** <https://orcid.org/0000-0003-2355-9029>

*** <https://orcid.org/0000-0001-5194-1714>

Napomena: Rad je deo projekta Ministarstva nauke, tehnološkog razvoja i inovacija, pod evidencionim brojem 451-03-137/2025-03/200096

Uvod

Forenzika se definiše kao skup naučnih znanja i tehničkih metoda koje se koriste za rešavanje slučajeva prvenstveno iz oblasti krivičnog prava, ali je prisutna i u okviru građanskog i upravnog prava (Tistarelli et al., 2014). Forenzička obrada materijala podrazumeva upotrebu naučnih metoda i tehnika za analiziranje fizičkih dokaza koji se koriste u policijskim istragama i sudskim procesima. Forenzika ima ključnu ulogu u rešavanju zločina (Kamble & Devarmani, 2024) jer predstavlja centralni segment kriminalističkih istraga i uključuje otkrivanje, sakupljanje i dokumentovanje dokaznog materijala, kao i njegovu analizu i međusobnu komparaciju. Posao forenzičara veštaka je da obezbedi dokaze koji su relevantni za potrebe sudskog procesa (Alam & Altiparmak, 2024). Stručno vođene kriminalističke istrage, pouzdane i precizna forenzička ispitivanja, kao i nepristrasan pravosudni sistem, predstavljaju osnovu uspešnog krivičnog pravosuđa (Jadhav et al., 2020).

Forenzika glasa i govora podrazumeva primenu rezultata naučnih istraživanja iz oblasti glasa i govora na slučajeve u pravnim procesima (Jain et al., 2024a). Preciznije, odnosi se na primenu metoda i tehnika koje se koriste za analizu glasa i govora u forenzičke svrhe (Šešum, 2021). Ona omogućava identifikaciju govornika, utvrđivanje autentičnosti snimaka, kao i analizu spornih izjava (Jain et al., 2024a). Najveći značaj forenzika glasa i govora ostvaruje u segmentu identifikacije govornika (Šešum, 2022). Biometrija omogućava pouzdanu i sistematičnu identifikaciju putem merenja i analiziranja bioloških podataka, a biometrija glasa i govora posebno je kompleksna zbog činjenice da zavisi od brojnih faktora (Jain & Kumar, 2010). Forenzička identifikacija govornika podrazumeva utvrđivanje njegovog identiteta na osnovu poređenja poznatog sa nepoznatim uzorkom govora u cilju detektovanja međusobnih sličnosti i razlika (Šešum, 2021). Svrha identifikacije govornika je, dakle, da se utvrdi da li glasovi koji predstavljaju dokazni materijal pripadaju istom govorniku ili različitim govornicima (Morrison et al., 2020).

Govor svake osobe, pored osnovne lingvističke poruke, sadrži i veliki broj informacija o samom govorniku, poput pola, starosti, porekla, zdravstvenog i emotivnog stanja itd. (Kovačević i Šešum, 2012). Identifikovanje govornika podrazumeva prepoznavanje određenih karakteristika glasa i govora na osnovu slušanja i pomoću specijalizovane računarske opreme (Patil & Parhi, 2009). Ovaj postupak posebno otežava to što govorna produkcija traje u vremenu i što su tokom nje sve karakteristike potencijalno promenjive (Šešum, 2021), zbog čega se forenzika glasa i govora smatra najkompleksnijom forenzičkom disciplinom (Šešum i Kovačević, 2015). Brojnost i promenljivost karakteristika glasa i govora nalažu potrebu za kombinovanjem lingvističkih analiza sa akustičkim analizama u cilju obezbeđivanja maksimalne pouzdanosti rezultata identifikacije govornika (Susanto & Nanda, 2024). Karakteristike glasa i govora načelno se dele na urođene i stečene, od kojih su prve uslovljene

anatomijom i fiziologijom govornog trakta, a druge govorom kojem smo bili izloženi tokom života (Šešum, 2013).

Pored promenljivosti parametara glasa i govora, identifikacija govornika je veoma kompleksan zadatak (Bonastre et al., 2015) i zbog toga što podrazumeva i druge vrste izazova, počevši od pravnog aspekta snimaka, preko njihovih tehničkih karakteristika, lingvističkog sadržaja itd. (Ajili, 2017). Zbog toga su i zahtevi forenzike glasa i govora složeni i vrlo precizni i odnose se i na kvalitet snimaka (tehničke i ambijentalne okolnosti) i na njihov kvantitet u pogledu trajanja govora osobe od interesa (Negrão & Domingues, 2021). S obzirom na to da se snimci mogu razlikovati u velikom broju aspekata, moglo bi se reći da svaki slučaj ima posebnu kombinaciju karakteristika. Neke od njih su ambijentalne prirode, npr. da li je snimak telefonski ili mikrofonski, zatim da li je i koliko prisutna pozadinska buka ili odjek, a druge se odnose na samog govornika, poput njegovog jezika, zdravstvenog stanja, intenziteta govora itd. (Durán et al., 2024). Većina snimaka koji predstavljaju predmet forenzičkih ispitivanja, očekivano, nije načinjena u okolnostima koje su optimalne, a često ni zadovoljavajuće i u pogledu ambijenta i u pogledu tehničkog aspekta snimanja (Šešum & Petrović, 2024).

Forenzika glasa i govora predstavlja veoma perspektivnu forenzičku disciplinu, za čijom primenom u savremenom svetu konstantno raste potreba (Šešum, 2021). Na međunarodnom nivou identifikacija govornika ima dugu tradiciju u kontekstu krivičnih istraga, a razvoj savremene tehnologije značajno je doprineo unapređenju ove metode; npr. pomogao je da se prevaziđu mnoge teškoće prilikom analize glasa i govora, uključujući i kvalitet govornog zapisa i višejezično prepoznavanje (Jain et al., 2024a).

Najaktuelnija nedoumica u oblasti forenzike glasa i govora odnosi se na to kako će razvoj tehnologije veštačke inteligencije uticati na budućnost ove forenzičke discipline. Iako je činjenica da trenutno nije moguće pružiti pouzdane odgovore na ovo pitanje zbog neverovatne brzine kojom se veštačka inteligencija još uvek razvija, veoma je važno predstaviti trenutne mogućnosti ovog kompjuterskog sistema, kao i iskustva korisnika sa njegovom dosadašnjom primenom u forenzici glasa i govora. Sistematično sagledavanje činjenica izloženih u literaturi treba da doprinese pripremi stručnjaka za adekvatno korišćenje ove tehnologije, a time i sprečavanju neželjenih ishoda do kojih bi njena neprimerena upotreba mogla dovesti.

Cilj

Cilj rada je da se pomoću sistematizacije dostupne literature predstave mogućnosti veštačke inteligencije u kontekstu forenzike glasa i govora, zatim ograničenja, kao i potencijalne opasnosti od njene primene u sudskim postupcima.

Metoda

Literatura koja je korišćena u radu obezbeđena je putem pretrage elektronskih baza podataka kojima je na internetu omogućen otvoren pristup. Pregled literature bio je usmeren ka novijim naučnim radovima, pa su za članak iskorišćena 24 rada napisana u poslednjih pet godina, ali i 22 rada koja su nastala pre tog perioda; tačnije 11 radova objavljenih u poslednjih 10 godina i 11 radova objavljenih pre 2015. godine. Za pronalaženje adekvatne literature korišćene su srpske i engleske verzije sledećih ključnih reči: forenzika, biometrija, glas i govor, veštačka inteligencija, automatska identifikacija, prepoznavanje govornika, savremeni sistemi, mašinsko učenje, neuronske mreže, duboko učenje, sudski postupci. Nakon detaljne analize literature, pristupilo se sintezi relevantnih nalaza.

Rezultati sa diskusijom

Mogućnosti veštačke inteligencije

Veštačka inteligencija, kao segment napredne tehnologije, predstavlja kompjuterski sistem koji je programiran da odgovara na zahteve i potrebe korisnika (Kamble & Devarmani, 2024). Veštačka inteligencija podrazumeva sposobnost kompjutera da obavlja zadatke uspešnije nego ljudi. Najlakše se može objasniti kao postupak replikacije ili simulacije ljudske inteligencije na nivou mašine (Dunsin et al., 2022).

Mašinsko učenje, duboko učenje i neuronske mreže su koncepti koji su međusobno povezani i koji pripadaju domenu veštačke inteligencije. **Mašinsko učenje** je aspekt veštačke inteligencije koji se odnosi na učenje mašine ili kompjutera na osnovu prethodnih aktivnosti ili obrazaca u cilju predviđanja novih izlaznih vrednosti korišćenjem specifičnih algoritama. Ono se smatra tehnikom rešavanja problema i transformisanjem upravljanja rizicima. **Duboko učenje** je funkcija veštačke inteligencije koja je dizajnirana da se koristi za otkrivanje objekata, prevođenje jezika, prepoznavanje govora i kritičko odlučivanje, putem simuliranja funkcionisanja ljudskog mozga u pogledu obrade podataka. Ono može da uči bez nadzora ljudi i koristi se za otkrivanje pranja novca i drugih vrsta prevara (Dunsin et al., 2022). Duboko učenje je kategorija mašinskog učenja, a algoritmi dubokog učenja koji imitiraju način na koji funkcioniše ljudski mozak nazivaju se veštačkim neuronskim mrežama (Anirudh, 2019). **Veštačke neuronske mreže** (*Artificial Neural Networks – ANN*) su kompjuterski mehanizmi koji simuliraju strukturalne i funkcionalne aspekte neuronskih mreža koje čine deo biološkog nervnog sistema. One su idealne za situacije koje zahtevaju predviđanja, klasifikaciju ili kontrolu u dinamičkim i kompleksnim kompjuterskim okruženjima (Bitter et al., 2010). Veštačke neuronske mreže sadrže veštačke neurone koji, kada se povežu, mogu da uče i rešavaju probleme. Neuronske mreže koje imaju sposobnost

učenja, procesuiranja podataka, samoorganizacije i prilagođavanja mogu rešavati probleme koji istovremeno podrazumevaju uslovljenost, nepreciznost i dvosmislenost (Dilek et al., 2015).

Primarna svrha veštačke inteligencije je da rešava probleme koje ljudi ne mogu rešiti i da pruža uvide koji bi se lako mogli zanemariti ukoliko bi se podaci posmatrali pojedinačno (Dunsin et al., 2022). Veštačka inteligencija, koja se u početku nazivala i „mašinskom inteligencijom”, datira od jula 1956. godine, kada je predstavljena na letnjem projektu istraživača koledža Dartmut. Ona može biti opisana na dva načina: kao naučna disciplina čiji je cilj da otkrije suštinu inteligencije i razvije inteligentne mašine i kao naučna disciplina čiji je cilj da otkrije metode za rešavanje kompleksnih problema koji se ne mogu rešiti bez primene nekog inteligentnog sistema (npr. odlučivanja zasnovanog na velikoj količini podataka) (Dilek et al., 2015).

Veštačka inteligencija može rešavati zadatke koji podrazumevaju vizuelnu percepciju, prepoznavanje govora, kogniciju, odlučivanje, učenje na osnovu uputstva i rešavanje kompleksnih problema, ponekad čak značajno brže i sa manje grešaka u odnosu na ljude (Jadhav et al., 2020). Ova popularna tehnologija razvija se veoma brzo i već je prisutna u različitim oblastima, između ostalog i u forenzici, gde se primarno koristi u svrhe ubrzavanja procesa veštačenja putem obavljanja pripremnih zadataka (Kamble & Devarmani, 2024), ali ograničenu primenu može naći i u pojedinim forenzičkim ekspertizama (Jadhav et al., 2020).

Prednosti primene veštačke inteligencije u forenzici

Upotreba veštačke inteligencije nameće nove trendove u oblasti krivičnog pravosuđa čineći je nezaobilaznim pomoćnim sredstvom u savremenim istragama (Kamble & Devarmani, 2024). Napredak ove tehnologije duboko je transformisao različite naučne discipline, pri čemu su se forenzičke nauke ispostavile kao posebno pogodno tlo za njenu primenu. Područje forenzike značajno je unapređeno i automatizovano zahvaljujući veštačkoj inteligenciji. U forenzičkoj primeni, ona nije samo inovativno, praktično sredstvo već može doprineti i pomeranju granica u okviru kojih se kreću forenzička ispitivanja (Kassem & Lodhi, 2024). Zbog raznovrsnosti i složenosti poslova koje podrazumevaju različite forenzičke discipline, upotreba veštačke inteligencije u ovoj oblasti može biti veoma korisna (Mitchell, 2014).

Veštačka inteligencija primenu u forenzici nalazi u okviru različitih vrsta kriminalističkih istraga (Kamble & Devarmani, 2024). Na sadašnjem stepenu razvoja ona može biti korisna u gotovo svim oblastima forenzike, zahvaljujući svojim raznovrsnim mogućnostima, kao što su: analiza podataka, prepoznavanje obrazaca, obrada slika, rudarenje podataka, statističke analize i probablističke metode, veštački mirisi, računarske i matematičke metode, grafičko modeliranje itd. (Jadhav et al., 2020). Smatra se da veštačka

inteligencija ima potencijal da transformiše savremenu forenziku, jer bi je mogla unaprediti ukoliko bude uspela da poboljša tačnost, efikasnost i brzinu dolaska do rezultata. Važnim doprinosom forenzici smatralo bi se i smanjenje mogućnosti ljudske greške, jer se tradicionalne metode često temelje na subjektivnim interpretacijama stručnjaka čiji nalazi zbog toga nekada mogu biti nedosledni (Arjamand et al., 2024). Potencijal veštačke inteligencije zasniva se i na činjenici da su uvidi koje generiše ovaj sistem daleko napredniji od onih zasnovanih na bilo kojoj drugoj dostupnoj tehnologiji (Dunsin et al., 2022).

Veštačka inteligencija može se koristiti za različite forenzičke svrhe, poput rekonstrukcije lica mesta, DNK analize, digitalne forenzike, kao i za mnoge druge namene (Arjamand et al., 2024). Proizvodi koji su zasnovani na ovoj tehnologiji mogu pomoći veštacima i istražiteljima da uštede vreme koje je potrebno za određene zadatke u različitim fazama istrage i ispitivanja. Ušteda vremena doprinosi efikasnijem radu policije, a time i pravosuđu koje je dostupnije građanima (Jadhav et al., 2020). Činjenica je da korišćenje veštačke inteligencije pomaže u tome da se veliki broj složenih podataka analizira za kratko vreme (Dunsin et al., 2022). Zahvaljujući veštačkoj inteligenciji moguće je obrađivati velike baze podataka i sprovoditi složene analize koje mogu služiti veštacima za potvrdu ili preispitivanje nalaza do kojih su sami došli. Na primer, moguće je brže pretraživanje baza DNK materijala ili otisaka prstiju, a time i ubrzavanje procesa veštačenja i celokupne kriminalističke istrage (Arjamand et al., 2024). Upotreba veštačke inteligencije u kontekstu forenzike, dakle, omogućava da se efikasno analiziraju različiti aspekti koje koriste standardni algoritmi da bi značajni podaci mogli biti brzo označeni, a uvidi svrstani u određenu kategoriju (Hoelz, 2009). Hibridni sistemi koji kombinuju inteligenciju ljudi i mašina takođe mogu biti korisni u kriminalističkim istragama (Franke & Srihari, 2007). S obzirom na to da tumačenje složenih situacija zahteva procenu stručnjaka, u formiranju konačnog mišljenja može učestvovati veštačka inteligencija dokle god poslednju reč ima stručno lice, tj. veštak (Kamble & Devarmani, 2024).

Primena veštačke inteligencije u forenzici glasa i govora

Napredak veštačke inteligencije uticao je na razvoj i funkcionisanje biometrijskih sistema (Jain et al., 2024b). U kontekstu biometrijske identifikacije, veštačka inteligencija doprinela je prepoznavanju osoba na osnovu lica, otisaka prstiju i glasa, jer je ove procese ubrzala i učinila pouzdanijim (Kassem & Lodhi, 2024). Biometrija glasa je poslednjih godina veoma napredovala zahvaljujući razvoju veštačke inteligencije, što je povećalo njenu bezbednost i prilagodljivost. Evolucija od tradicionalnih metoda prepoznavanja govornika do novih tehnika koje se zasnivaju na upotrebi veštačke inteligencije bila je praćena mnogim izazovima, kao što su: stabilnost glasa, multijezičko prepoznavanje i otpornost

na opstrukcije i lažiranja. Međutim, pojava sofisticiranih metoda lažiranja i kloniranih snimaka nameće potrebu za stalnim inovacijama tehnologije kako bi se razvile pouzdanije metode za otkrivanje falsifikata i unapredila otpornost sistema za biometrijsku analizu glasa i govora (Jain et al., 2024b).

Kao i u slučaju drugih forenzičkih disciplina, veštačka inteligencija mogla bi da ima značajnu, čak i revolucionarnu ulogu u forenzici glasa i govora ukoliko bude omogućila sprovođenje brzih i preciznih analiza govornih uzoraka u cilju prepoznavanja govornika i otkrivanja maskiranja glasa i govora (Jain et al., 2024a), što predstavlja namerno manipulisanje glasom i govorom da bi se promenile njihove karakteristike i time onemogućila identifikacija govornika (Šešum, 2024). Na primer, korisna upotreba mogućnosti veštačke inteligencije odnosila bi se na praćenja telefonskih poziva i poređenje uzoraka govora sa bazom snimaka u cilju prepoznavanja glasa osumnjičene osobe ili žrtve. Takođe, i obezbeđivanje mogućnosti za otkrivanje maskiranog glasa i govora putem prepoznavanja njihovih osnovnih obrazaca koji mogu ukazati na identitet govornika predstavljalo bi veoma značajan pomak (Phonexia launches voice biometrics product for forensic investigations, 2024). Za sada se sistemi veštačke inteligencije smatraju značajnim pomoćnim sredstvima koja mogu naći ograničenu primenu u analizi glasa i govora (Jain et al., 2024a).

Glas i govor su jedinstveni, baš kao i otisak prsta i DNK materijal, sa tom razlikom što su, pored unutrašnjih, anatomskih i fizioloških, uslovljeni i spoljašnjim, društvenim faktorima. Programi koji se baziraju na veštačkoj inteligenciji koriste ovu jedinstvenost za potvrđivanje identiteta, analizirajući glas i govor putem algoritama mašinskog učenja koji su usmereni na karakteristike poput osnovne frekvencije, intonacije i akcenta (Jain et al., 2024a). Integracija mašinskog učenja u tehniku biometrijskog prepoznavanja glasa i govora fokusira se na efikasno izdvajanje karakteristika, što značajno poboljšava prepoznavanje (Jain et al., 2024b). Izolacija karakteristika je ključna za tačnost postignuća sistema za identifikaciju govornika i njegovo biometrijsko prepoznavanje (Jain et al., 2024b). Modeli mašinskog učenja su na skupu uzoraka uvežbavani da prepoznaju govorne obrasce, ton i druge karakteristike govornika. Ova tehnologija može biti posebno korisna u slučajevima koji uključuju pretnje, prevare kao i bilo koje druge kriminalne aktivnosti gde se glas i govor pojavljuju kao primarni dokazi. Sofisticiranost tehnologije veštačke inteligencije u ovoj oblasti omogućava razlikovanje pojedinaca čak i kada je glas oštećen ili izobličen (Jansen et al., 2021; Singh et al., 2018).

Pošto je potencijalna količina dostupnih podataka ogromna, primena dubokog učenja može biti najbolji način za njihovu klasifikaciju (Al Ajmi et al., 2022), jer algoritmi dubokog učenja mogu da skeniraju karakteristike glasa i govora govornika u cilju kreiranja njegovog ličnog biometrijskog otiska (Hong & Cho, 2021). Ovi otisci se mogu koristiti za identifikaciju osumnjičenih ili

žrtava u okviru bilo kog prestupa koji podrazumeva glas i govor kao dokazni materijal (Arjamand et al., 2024). Kada postoji potreba za utvrđivanjem autentičnosti glasa i sa zvučnih snimaka, upotrebljavaju se algoritmi veštačke inteligencije koji se zasnivaju na neuronskim mrežama (Al Ajmi et al., 2022). Pored navedenih funkcija, veštačka inteligencija može doprineti i poboljšanju zvučnih snimaka putem eliminisanja pozadinske buke, što će uticati na razumljivost izgovorenog sadržaja (Arjamand et al., 2024).

Iako integracija različitih koncepata veštačke inteligencije, kao što su mašinsko učenje, duboko učenje i neuralne mreže, pokazuje veliki potencijal za poboljšanje otpornosti i performansi sistema, izazovi poput varijabilnosti glasa i govora, narušavanja privatnosti govornika i obima forenzičke primene ovih tehnologija i dalje nisu prevaziđeni (Jain et al., 2024b). Na primer, isti govornik iste glasove neretko izgovara na različite načine u zavisnosti od njihove pozicije unutar reči ili u dodiru dveju reči, kao i u zavisnosti od neposrednog glasovnog okruženja, tj. konteksta, pa će neakcentovani vokali, na primer, biti drugačije produkovani, a samim tim i imati drugačiju akustičku sliku nego akcentovani (Ivanović i Šešum, 2009). Takođe, forenzička ispitivanja glasa i govora, između ostalog, karakterišu i složenost biometrijskih performansi u višejezičkom kontekstu, a posebno u kontekstu identifikacije govornika, što predstavlja veliki izazov sistemu koji nameće jezička raznolikost (de Lima & Da-Costa Abreu, 2022). Iako veštačka inteligencija još uvek nije dorasla primeni u stvarnim forenzičkim slučajevima, činjenica je da predstavlja jaku osnovu za budućnost i da bi mogla da pruži novu perspektivu narednim generacijama veštaka za glas i govor (Jain et al., 2024a).

Ipak, pomoć veštačke inteligencije u procesu forenzičkih ispitivanja može biti od velikog značaja jer se najrealnija potreba za automatizacijom procesa forenzičke ekspertize glasa i govora odnosi na otkrivanje govora na snimcima i na njegovu transkripciju, budući da su to dve pripremne etape veštačenja koje oduzimaju dosta vremena, a ne zahtevaju specijalizovano znanje veštaka. Detekcija govora, poznata i kao detekcija glasovne aktivnosti (*Voice Activity Detection* – VAD) podrazumeva uočavanje govornih deonica na audio-snimcima. Preciznost programa koji tome služe je veoma važna jer, ukoliko nisu precizni, oni mogu označiti pozadinsku buku kao govor, ili obratno, što može dovesti do konfuzije u postupku veštačenja čije bi razrešavanje verovatno oduzelo više vremena nego što bi se primenom ovih programa uštedelo (Negrão & Domingues, 2021). Transkripcija govora podrazumeva njegovo prenošenje u tekst. Kompjuterski programi koji obavljaju transkripciju govora su već dugo prisutni na tržištu, a njihove novije verzije uglavnom su zasnovane na modelima veštačke inteligencije.

Automatska identifikacija govornika

Savremeni trend u forenzici glasa i govora predstavlja orijentacija ka automatizaciji procesa veštačenja (Šešum, 2022). Automatska identifikacija govornika počela je da privlači značajniju pažnju stručne javnosti tokom prve decenije 21. veka, kada se dogodio i njen nagli razvoj (Šešum, 2021). Ova metoda identifikacije govornika zasniva se na programiranju automatskog sistema da sam prepozna ljudski govor (Šešum & Petrović, 2024). Zbog potencijalne jezičke zabune važno je razjasniti terminologiju koja je u upotrebi. Automatsko prepoznavanje govora odnosi se na prepoznavanje rečenog u cilju izvršavanja komandi ili formiranja transkripta, a automatsko prepoznavanje govornika na prepoznavanje govornika na osnovu njegovog glasa i govora (Singh, 2019). Sistemi koji se koriste za prepoznavanje govornika posebno su kreirani za ovu namenu, a mogu se zasnivati na tehnologiji veštačke inteligencije (npr. *Batvox*), ali i ne moraju (npr. *Alize*), iako se potonji smatraju donekle prevaziđenima.

Sistemi za automatsko prepoznavanje govornika imaju širok spektar primene u brojnim oblastima, poput bezbednosnih sistema, personalizovanih asistivnih servisa i forenzičkih istraga (Ben Amor, 2024). U forenzičkoj primeni svrha ove tehnologije je da se utvrdi da li dva uzorka govora pripadaju istom govorniku, a funkcioniše tako što automatski dobijeni rezultati pokazuju na koji od uzoraka pohranjenih u dostupnoj bazi podataka najviše liči sporni govor. Problem sa ovakvim načinom dobijanja rezultata je taj što postoji ozbiljan nivo verovatnoće da će softver u bazi svakako pronaći govornika koji najviše odgovara spornom uzorku, iako se govornik za kojim se traga možda uopšte ni ne nalazi u bazi (Šešum, 2021). Automatski sistemi se, dakle, oslanjaju na obrasce verovatnoće pojave određenih karakteristika govora umesto na njihovu pojedinačnu analizu (Šešum & Petrović, 2024). Kada se zasnivaju na veštačkoj inteligenciji, primarno su utemeljeni na kompleksnim dubokim neuronskim mrežama i svoje zaključke predstavljaju u vidu jedne vrednosti (Ben Amor, 2024). Korišćenje dubokog učenja može značajno unaprediti kvalitet automatske identifikacije govornika ukoliko se dostigne potreban nivo tačnosti ovih modela (Watanabe et al., 2017).

Ipak, ovu metodu još uvek karakterišu mnoga ograničenja. Na primer, tehnologija automatske identifikacije govornika je dosta osetljiva, čak i u veoma dobrim uslovima, jer svaka iole značajnija promena glasa i govora može dovesti do neprepoznavanja iako je u pitanju isti govornik (Singh, 2019). Takođe, većina automatskih sistema oslanja se isključivo na biološke, uslovljene karakteristike govora, sasvim zanemarujući podjednako važne stečene govorne navike (Šešum & Petrović, 2024). Zbog njihove zatvorenosti automatski sistemi identifikacije govornika često se nazivaju crnim kutijama. Uprkos određenim prednostima koje se uglavnom odnose na brzinu pretraga baza, oni ne pružaju zadovoljavajući nivo transparentnosti i objašnjivosti postupaka koje primenjuju te rezultata koje prezentuju (Ben Amor, 2024).

Nedovoljna transparentnost ovih sistema predstavlja njihovu slabu tačku u kontekstu njihove pouzdanosti (Hall et al., 2021), jer je pouzdanost sistema usko povezana sa transparentnošću. Zbog netransparentnosti upotreba automatske identifikacije govornika nije prihvatljiva na sudu jer nalaze koje ovaj sistem generiše nije moguće prilagoditi i objasniti u skladu sa standardima pravosuđa (Mitchell, 2014). Takođe, automatska identifikacija govornika gotovo u potpunosti isključuje faktor ljudskog sagledavanja i zaključivanja, što je naučno neopravdano s obzirom na činjenicu da je reč o analizi realizacije jezika kao najsloženijeg društvenog fenomena (Šešum, 2021). Može se reći da je u slučaju automatske identifikacije uloga stručnjaka svedena isključivo na operatera programa (Šešum, 2022).

Iako primena veštačke inteligencije može biti ključna u različitim forenzičkim disciplinama (Jadhav et al., 2020), forenzika glasa i govora nije jedna od njih. Automatski sistemi za identifikaciju govora nisu se pokazali dovoljno efikasnim, a pritom su vrlo zahtevni budući da podrazumevaju veliku bazu reči i fraza, kao i konstantno usavršavanje jezičkih i akustičkih modela (Endah et al., 2017). Zbog toga se veštaci koji rade u forenzičkim centrima radije oslanjaju na tradicionalnu tehnologiju koju primenjuju za sprovođenje preciznih i tačnih analiza, a upotreba veštačke inteligencije olakšava im rad utoliko što može da ubrza određene etape postupka (Kamble & Devarmani, 2024). Metodom identifikacije govornika, koja je u potpunosti pouzdana, smatra se samo tradicionalna, auditivno-instrumentalna metoda, koja predstavlja kombinaciju slušanja od strane stručnjaka i analize glasa i govora uz pomoć kompjuterskog softvera (Šešum, 2021). Iako su sistemi za automatsko prepoznavanje govornika postigli značajan napredak u pogledu tačnosti nalaza (Durán et al., 2024), njihova primena u realnim forenzičkim situacijama još uvek se izbegava. S obzirom na kompleksnost glasa i govora razumljivo je zašto pokušaji potpune automatizacije procesa identifikacije govornika, iako traju već godinama, do sada nisu bili uspešni (Šešum, 2022). Ipak, izvesno je da će forenzika glasa i govora, kao i sve druge forenzičke discipline, nastaviti da se razvija u smeru što veće automatizacije radnog procesa.

Etičke nedoumice u primeni veštačke inteligencije

Sa napretkom novih tehnologija nameće se i potreba za temeljnim razmatranjem etičkih pitanja koja se odnose na njihovu primenu (Kassem & Lodhi, 2024). Na primer, u oblasti forenzike glasa i govora od početka primene automatskih sistema postavlja se pitanje da li i kada treba dopustiti uključivanje takvih algoritamskih „crnih kutija” u krivične istrage, što je pitanje koje, pored naučnog, ima i etički, pravni i filozofski značaj (Durán et al., 2024). Nedostatak transparentnosti onemogućava ljude da razumeju kako se sistem ponaša i na temelju čega donosi konkretne odluke. Sistem čije se odluke ne mogu objasniti, a posledično ni proveriti, nije sistem kome je lako verovati, posebno kada je

u pitanju njegova primena u osetljivim područjima poput zdravstvene zaštite ili forenzike, gde je pridržavanje normi morala i etike od najveće važnosti (Ben Amor, 2024). Sistemi koji se koriste u forenzičkim naukama moraju biti objašnjivi, tj. mora biti jasno kako su i na osnovu čega doneseni konkretni zaključci i predikcije. To je neophodno da bi se obezbedilo poverenje policije, pravosuđa i javnosti u nalaze ovih sistema (Kassem & Lodhi, 2024).

Integracija sistema veštačke inteligencije u različite oblasti, uključujući i forenzičke nauke, potakla je značajne debate u pogledu etičnosti njihove primene i dovela do zabrinutosti u vezi sa privatnošću, bezbednošću, pristrasnošću i pouzdanošću (Kassem & Lodhi, 2024). Etičke dileme u pogledu upotrebe veštačke inteligencije u forenzičke svrhe veoma su prisutne, što nije iznenađujuće s obzirom na činjenicu da korišćenje ove tehnologije može imati veliki uticaj na živote ljudi (Durán et al., 2024). Da li treba verovati veštačkoj inteligenciji aktuelno je pitanje kako u krugovima stručnjaka, tako i u široj društvenoj zajednici (Franke, 2023). Jedna od najvažnijih etičkih nedoumica u pogledu korišćenja ovih sistema u forenzici odnosi se na uticaj koji ovakve tehnologije mogu imati na sudske procese jer se veštačka inteligencija može koristiti za analiziranje kompleksnih baza podataka i pružanja uvida koji mogu uticati na ishode sudskih procesa. Problemi sa pouzdanošću tehnologija mogu imati veoma negativne posledice u vidu pogrešnih odluka suda. Stoga je osiguranje tačnosti, transparentnosti i nepristrasnosti analiza koje pokreće veštačka inteligencija od krucijalnog značaja (Slovak & Shepherd, 2020).

U kriminalističkim istragama veštaci daju mišljenje o vrednosti dokaza, koje se zasniva na njihovoj ekspertizi. Poverenje sudija u mišljenje veštaka temelji se na kredibilitnosti veštaka, tj. na činjenici da su stručni i nepristrasni. Na sudu veštaci razjašnjavaju svoje nalaze i navode koje su karakteristike dokaza i zbog čega uzimali u obzir prilikom procenjivanja njihovog značaja, kako bi dodatno doprineli poverenju suda u njihove nalaze (Durán et al., 2024). Detaljno objašnjavanje nalaza i mišljenja veštaka potrebno je jer je veštačenje kompleksan postupak, a značaj dokaza ključan u kontekstu sudskog procesa. Veštaci su dužni da, po potrebi, objasne i kriterijume, zatim metodologiju koja je prethodila davanju mišljenja, kao i različite faktore koji su doprineli njegovom formiranju (O'Hear, 2019). Nasuprot tome, sistemi veštačke inteligencije nisu objašnjivi u klasičnom smislu, tj. stručnjaci ne mogu reći na šta se odnose njihovi pojedinačni parametri niti predvideti koji bi bio ishod primene sistema ukoliko bi se podaci kojima raspolažu prema nekom kriterijumu razlikovali (Durán et al., 2024). Stoga se kao imperativ poboljšanja pouzdanosti sistema veštačke inteligencije nameće unapređenje njihove transparentnosti i objašnjivosti (Ben Amor, 2024).

S obzirom na to da tačnost i pouzdanost primene sistema veštačke inteligencije zavisi, između ostalog, i od kvaliteta podataka na kojima su uvežbavani, uvidi koje su generisali mogu biti narušeni greškama ili pogrešnim

tumačenjima, što može voditi ka pogrešnom usmeravanju istrage (Kamble & Devarmani, 2024). Modeli veštačke inteligencije zasnovani su na obrascima koje algoritam pronalazi u velikoj količini podataka, za koje nema garancije da su zaista relevantni. Jedan od primera pogrešnog zaključivanja predstavlja iskustvo u primeni modela koji razlikuje pse haskije od vukova (Ribeiro et al., 2016), gde je veštačka inteligencija zasnovala razlikovanje na osnovu prisustva snega na fotografijama, što je dovelo do njenog uverenja da samo vukovi mogu biti na snegu i da im je to razlikovno obeležje u odnosu na haskije (Durán et al., 2024).

Pronalaženje efikasnog načina za iskorištavanje prednosti veštačke inteligencije, uz istovremeno poštovanje etičkih standarda i privatnosti ljudi, još uvek predstavlja veliki izazov (Kamble & Devarmani, 2024). Iako ova tehnologija nudi značajne mogućnosti za napredak forenzike, takođe donosi i niz izazova koje je neophodno prevazići pre njenog uključivanja u realne sudske procese. Prevazilaženje ovih izazova podrazumeva kombinaciju unapređenih tehnoloških rešenja, zakonskog nadzora i etičkih preporuka (Kassem & Lodhi, 2024).

Zaključak

Veštačka inteligencija, iako predstavlja najveće dostignuće savremene tehnologije, u kontekstu forenzike glasa i govora još uvek ima veoma ograničenu primenu, koja se pretežno odnosi na ubrzavanje procesa veštačenja putem detekcije govora zabeleženog na zvučnim snimcima, kao i njegovog pretvaranja iz zvučnog u tekstovni oblik. Iako dosta efikasni, sistemi veštačke inteligencije koji su konstruisani za potrebe identifikacije govornika ne pružaju dovoljnu transparentnost, koja je neophodan uslov za njihovo korišćenje u sudskim postupcima.

U cilju prihvatanja automatske identifikacije govornika kao legitimne forenzičke metode, usavršavanje ovog sistema je neophodno usmeravati ka poboljšanju preciznosti i transparentnosti rada, kao i ka unapređivanju otpornosti sistema na odstupanja govornih uzoraka koji su sadržani u bazama. S obzirom na spektar mogućnosti veštačke inteligencije koje se neprekidno uvećavaju, izvesno je da će budućnost mnogih forenzičkih disciplina, pa i forenzike glasa i govora u velikoj meri biti uslovljena daljim razvojem ovih pametnih sistema. Jedan od značajnijih izazova koje je na tom putu neophodno prevazići odnosi se na etičnost primene veštačke inteligencije u sudskim procesima.

Literatura

- Ajili, M. (2017). *Reliability of voice comparison for forensic applications* [Doctoral dissertation, Université d'Avignon et des Pays de Vaucluse]. HAL Open Science. <https://theses.hal.science/tel-01774394v1/file/pdf2star-1524052865-Manuscrit-de-these-AJILI-19-10-2017.pdf>
- Al Ajmi, S., Hayat, K., Al Obaidi, A. M., Kumar, N., Najmuldeen, M., & Magnier, B. (2022). Digital audio forensics: Blind human voice mimicry detection. *arXiv preprint arXiv:2209.12573*. <https://doi.org/10.48550/arXiv.2209.12573>
- Alam, S., & Altiparmak, Z. (2024). XAI-CF – Examining the role of explainable artificial intelligence in cyber forensics. *arXiv:2402.02452*. <https://doi.org/10.48550/arXiv.2402.02452>
- Anirudh, V. K. (2019). *What Is Deep Learning: Definition, Framework, and Neural Networks* [Web log post]. <https://www.toolbox.com/tech/artificial-intelligence/tech-101/what-is-deep-learning-definition-framework-and-neural-networks/>
- Arjamand, M., Saleem, A., Basit, A., Iftikhar, S., Sharif, M., Cholistani, M. S., Farhan, M., Shumail, S., Khan, B. A., Ali, Z., Shahid, B., & Hasnain, M. (2024). The role of artificial intelligence in forensic science: Transforming investigations through technology. *International Journal of Multidisciplinary Research and Publications (IJMRAP)*, 7(5), 67-70. <https://ijmrmap.com/wp-content/uploads/2024/10/IJMRAP-V7N5P52Y24.pdf>
- Ben Amor, I. (2024). *Deep modeling based on voice attributes for explainable speaker recognition: Application in the forensic domain* [Doctoral dissertation, Université d'Avignon]. HAL Archives. <https://hal.science/tel-04634215>
- Bitter, C., Elizondo, D. A., & Watson, T. (2010). Application of artificial neural networks and related techniques to intrusion detection. *IEEE World Congress on Computational Intelligence (WCCI 2010)*, 949-954. <https://doi.org/10.1109/IJCNN.2010.5596532>
- Bonastre, J. F., Campbell, J. P., & Mason, J. S. D. (2010). Forensic speaker recognition. *IEEE Signal Processing Magazine*, 27(2), 95-103. <https://doi.org/10.1109/MSP.2010.936730>
- de Lima, T. A., & Da-Costa Abreu, M. C. (2022). Phoneme analysis for multiple languages with fuzzy-based speaker identification. *IET Biometrics*, 11(6), 614-624. <https://doi.org/10.1049/bme2.12078>
- Dilek, S., Çakır, H., & Aydın, M. (2015). Applications of artificial intelligence techniques to combating cyber crimes: A review. *International Journal of Artificial Intelligence & Applications (IJAIA)*, 6(1), 21-39. <https://doi.org/10.5121/ijaia.2015.6102>
- Dunsin, D., Ghanem, M. C., & Ouazzane, K. (2022). The use of artificial intelligence in digital forensics and incident response in a constrained environment. *International Journal of Information and Communication Engineering*, 16(8), 280-285.
- Durán, J. M., van der Vloed, D., Ruifrok, A., & Ypma, R. J. F. (2024). From understanding to justifying: Computational reliabilism for AI-based forensic evidence evaluation. *Forensic Science International: Synergy*, 9, 100554. <https://doi.org/10.1016/j.fsisyn.2024.100554>
- Endah, S. N., Adhy, S., & Sutikno, S. (2017). Comparison of feature extraction MFCC and LPC in automatic speech recognition for Indonesian. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, 15(1), 292-298. <https://doi.org/10.12928/telkomnika.v15i1.3605>
- Franke, K. (2023). *Artificial intelligence (AI) in forensic sciences*. Wiley-Blackwell.
- Franke, K., & Srihari, S. N. (2007). Computational forensics: Towards hybrid-intelligent crime investigation. In *Third International Symposium on Information Assurance and Security*, (pp. 383-386), Manchester. doi: 10.1109/IAS.2007.84.

- Hall, S. W., Sakzad, A., & Choo, K. R. (2021). Explainable artificial intelligence for digital forensics. *WIREs Forensic Science*, 4(2), 1-11. <https://doi.org/10.1002/wfs2.1434>
- Hoelz, B., Ralha, C., & Geeverghese, R. (2009). Artificial intelligence applied to computer forensics. U *Proceedings of the 2009 ACM Symposium on Applied Computing – SAC '09* (pp. 1181-1185). <https://doi.org/10.1145/1529282.1529549>
- Ivanović, M., & Šešum, M. (2009). *Jedan tip regionalne redukcije neakcentovanih slogova kao forenzički marker*. U *Zbornik radova 53. konferencije ETRAN* (str. 1-4). Vrnjačka Banja, Srbija: Društvo za ETRAN.
- Jadhav, E. B., Sankhla, M. S., & Kumar, R. (2020). Artificial intelligence: Advancing automation in forensic science & criminal investigation. *Journal of Seybold Report*, 15(8), 2064-2075. <https://doi.org/10.2174/2666484401666220819111603>
- Jain, A. K., & Kumar, A. (2010). Biometrics of next generation: An overview. *Second Generation Biometrics*, 12(1), 2-3.
- Jain, P., Chinmayee, P., Kaur, K., Chaudhary, S., Kaur, K., & Karunya, S. (2024a). Advancements in forensic voice analysis: Legal frameworks and technology integration. *Asian Journal of Advances in Research*, 7(1), 369-384.
- Jain, P., Ujjainia, P., Srivastava, A., Shrivastav, K., Rani, I., Vashisht, A., Behera, R., Moza, B., & Mukherjee, D. (2024b). Forensic perspective on voice biometrics and AI: A review. *International Journal of Scientific Research in Science and Technology*, 11(5), 49-63. <https://doi.org/10.32628/IJSRST2411581>
- Jansen, F., Sánchez-Monedero, J., & Dencik, L. (2021). Biometric identity systems in law enforcement and the politics of (voice) recognition: The case of SiiP. *Big Data & Society*, 8(2), 49-63. <https://doi.org/10.1177/20539517211063604>
- Kamble, Y., & Devarmani, N. G. (2024). Application of artificial intelligence (AI) in forensic investigation. In R. Mia, V. Panchal, S. Singla, & M. Jain (Eds.), *Advanced and emerging development in multidisciplinary approaches in forensic science, digital forensic, law and criminology* (pp. 197-204). AkiNik Publications. <https://doi.org/10.22271/ed.book.2825>
- Kassem, M. A., & Lodhi, K. (2024). Revolutionizing forensic science: The role of artificial intelligence and machine learning. *International Journal of Data Science*, vi, 10-18. <https://doi.org/10.5147/ijds.vi.255>
- Kovačević, J., & Šešum, M. (2012). *Uticaj mehaničke modifikacije govora na akustičku strukturu govornih segmenata*. U *Zbornik radova 56. konferencije ETRAN* (str. 1-4). Zlatibor, Srbija: Društvo za ETRAN.
- Mitchell, F. (2014). The use of artificial intelligence in digital forensics: An introduction. *Digital Evidence and Electronic Signature Law Review*, 7. <https://doi.org/10.14296/deeslr.v7i0.1922>
- Morrison, G. S., Enzinger, E., Ramos, D., González-Rodríguez, J., & Lozano-Díez, A. (2020). Statistical models in forensic voice comparison. In D. L. Banks, K. Kafadar, D. H. Kaye, & M. Tackett (Eds.), *Handbook of forensic statistics* (Ch. 20, pp. 451-497). CRC Press.
- Negrão, M., & Domingues, P. (2021). SpeechToText: An open-source software for automatic detection and transcription of voice recordings in digital forensics. *Forensic Science International: Digital Investigation*, 38, 301223. <https://doi.org/10.1016/j.fsidi.2021.301223>
- O'Hear, M. M. (2019). Appellate review of sentence explanations: Learning from the Wisconsin and federal experiences. *Marquette Law Review*, 103(2), 641-678. https://scholarship.law.marquette.edu/mulr/vol93/iss2/15/?utm_source=chatgpt.com

- Patil, H. A., & Parhi, K. K. (2009). Variable length Teager energy based Mel cepstral features for identification of twins. In S. Chaudhury, A. N. Rajagopalan, D. S. Chauhan, S. Murty, & P. K. Biswas (Eds.), *Pattern Recognition and Machine Intelligence* (pp. 525-530). Springer. https://doi.org/10.1007/978-3-642-11164-8_85
- Phonexia launches voice biometrics product for forensic investigations. (2024) (Web stranica). <https://www.biometricupdate.com/202401/phonexia-launches-voice-biometrics-product-for-forensic-investigations>
- Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?": Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1135-1144). ACM. <https://doi.org/10.1145/2939672.2939778>
- Singh, N., Agrawal, A., & Khan, R. A. (2018). Voice biometric: A technology for voice-based authentication. *Advanced Science, Engineering and Medicine*, 10(7-8), 754-759. <https://doi.org/10.1166/asem.2018.2219>
- Singh, S. (2019). The role of speech technology in biometrics, forensics and man-machine interface. *International Journal of Electrical and Computer Engineering*, 9(1), 281-288. <https://doi.org/10.11591/ijece.v9i1.pp281-288>
- Slovak, B. L., & Shepherd, S. M. (2020). Machine learning and forensic risk assessment: New frontiers. *The Journal of Forensic Psychiatry & Psychology*, 31(4), 571-581. <https://doi.org/10.1080/14789949.2020.1779783>
- Susanto, S., & Nanda, D. S. (2024). Predicting the classification of high vowel sound by using artificial neural network: A study in forensic linguistics. *IAES International Journal of Artificial Intelligence (IJ-AI)*, 13(1), 195-200. <https://doi.org/10.11591/ijai.v13.i1.pp195-200>
- Šešum, M. (2013). Komparativna analiza formantnih struktura glasova sestara i glasova monozigotnih bliznakinja. *Beogradska defektološka škola*, 19(3), 515-527.
- Šešum, M. (2021). Forenzička fonetika-identifikacija govornika. U S. Knežević (Ur.), *Forenzičko računovodstvo, istražne radnje, ljudski faktor i primenjeni alati* (str. 828-859). Beograd: Fakultet organizacionih nauka.
- Šešum, M. (2022). Uloga surdologa i logopeda u forenzičkoj analizi glasa i govora. U *V simpozijum logopeda Srbije „Tinski rad u logopediji i defektologiji“* (str. 280-290). Udruženje logopeda Srbije.
- Šešum, M. (2024). Forenzički značaj maskiranja glasa i govora. U *Zbornik sažetaka radova IX simpozijuma logopeda Srbije: Transdisciplinarnost u logopediji danas* (str. 11). Udruženje logopeda Srbije.
- Šešum, M., i Kovačević, J. (2015). Forenzička fonetika. *Vodič za primenu Zakonika o krivičnom postupku Republike Srbije* (str. 90-105). Grafol.
- Šešum, M., & Petrović, E. (2024). The complexity of speaker identification based on voice and speech and its application in forensics. In S. H. Fazeli (Ed.), *The Ninth International Conference on Languages, Linguistics, Translation and Literature – Full Articles (Volume One)* (pp. 90-108), Ahwaz Publication of Research and Sciences.
- Tistarelli, M., Grosso, E., & Meuwly, D. (2014). Biometrics in forensic science: Challenges, lessons and new technologies. U *Proceedings of the International Workshop on Biometric Authentication* (pp. 153-164). Springer. https://doi.org/10.1007/978-3-319-05050-8_15
- Watanabe, S., Delcroix, M., Metze, F., & Hershey, J. R. (Eds.). (2017). *New era for robust speech recognition*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-64680-0>

Artificial intelligence in voice and speech forensics

Mia M. Šešum, Marina N. Šestić, Maja P. Ivanović

University in Belgrade – Faculty of Special Education and Rehabilitation, Belgrade, Serbia

Introduction. Artificial intelligence represents the pinnacle of modern technological development. Designed to assist humans in performing tasks that are burdensome or even exceed human capabilities, this computer-based system finds application in various areas of professional and everyday functioning. One of the fields where the resources provided by artificial intelligence can be of particular value is forensic science. *Objective.* The aim of this paper is to present, through a review of the available literature, the potential applications of artificial intelligence in the context of voice and speech forensics, its limitations, and the potential risks associated with its use in judicial proceedings. *Methods.* The literature used in this study was obtained by searching accessible databases via remote access, followed by a synthesis of relevant findings. *Results.* Artificial intelligence is a highly sophisticated technological advancement with the potential to transform various forensic disciplines in the future. Due to the fact that language is one of the most complex social phenomena, as well as the consequences that could arise from analyses whose inner workings are not transparent to expert witnesses, this technology has not yet been included among the officially recognized and accepted methods of voice forensic analysis. *Conclusion.* Artificial intelligence can contribute to forensic voice and speech examinations. However, given that the lack of transparency of such systems raises fundamental ethical concerns, its application remains, for the time being, limited to operations conducted during the preparatory phase of forensic analysis. Increasing the transparency of AI-based systems, along with ongoing efforts to improve their robustness and accuracy, will help resolve ethical dilemmas and support the acceptance of this technology as legitimate within judicial proceedings.

Keywords: forensics, voice and speech, artificial intelligence, automatic identification, biometry

PRIMLJENO: 15.4.2025.
PRIHVAĆENO: 02.9.2025.